



WORKSHEET

Data Representation Question – Box and Whisker Plots

Question Description: Every year in your Statistics 1 exam there is a question on Representing Data. Below are all the exam questions when they ask you to represent data by drawing a “Box and Whisker Plot”.

QUESTIONS

Q1 s2016_p63

A group of children played a computer game which measured their time in seconds to perform a certain task. A summary of the times taken by girls and boys in the group is shown below.

	Minimum	Lower quartile	Median	Upper quartile	Maximum
Girls	5	5.5	7	9	13
Boys	4	6	8.5	11	16

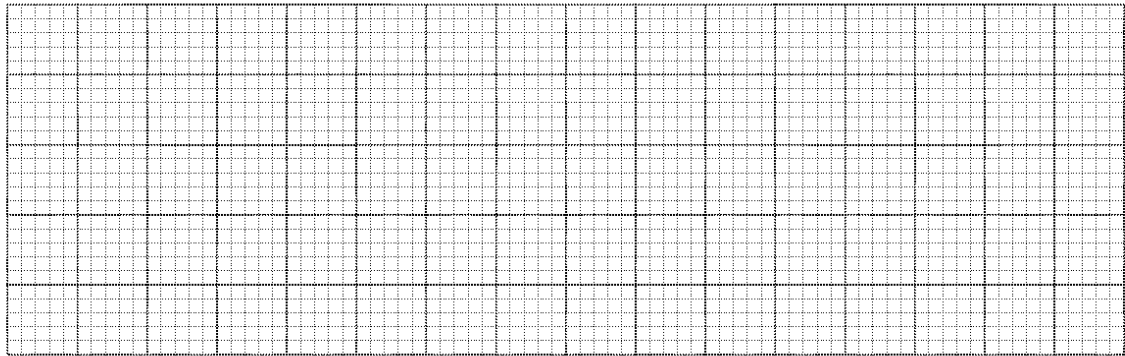
- (i) On graph paper, draw two box-and-whisker plots in a single diagram to illustrate the times taken by girls and boys to perform this task. [3]
- (ii) State two comparisons of the times taken by girls and boys. [2]

Q2 s2018_p61

In a survey 55 students were asked to record, to the nearest kilometre, the total number of kilometres they travelled to school in a particular week. The results are shown below.

5	5	9	10	13	13	13	15	15	15	15
16	18	18	18	19	19	20	20	20	20	21
21	21	21	23	25	25	27	27	29	30	33
35	38	39	40	42	45	48	50	50	51	51
52	55	57	57	60	61	64	65	66	69	70

- (i) On the grid, draw a box-and-whisker plot to illustrate the data. [5]



An 'outlier' is defined as any data value which is more than 1.5 times the interquartile range above the upper quartile, or more than 1.5 times the interquartile range below the lower quartile.

- (ii) Show that there are no outliers. [2]

Q3 w2018_p62

The following back-to-back stem-and-leaf diagram shows the reaction times in seconds in an experiment involving two groups of people, A and B .

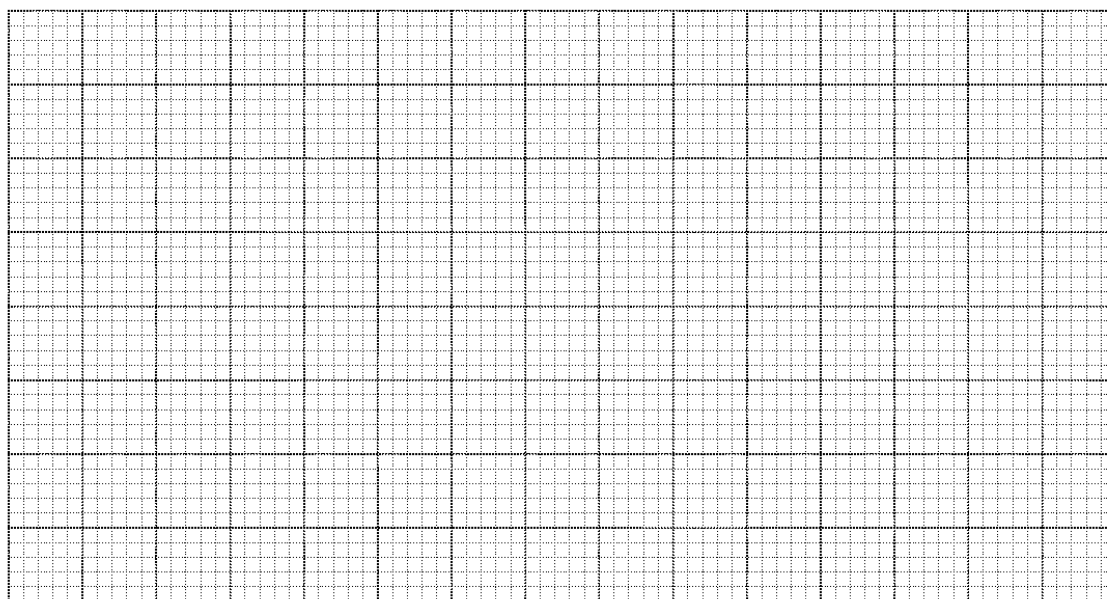
A			B	
(4)	4 2 0 0	20	5 6 7	(3)
(5)	9 8 5 0 0	21	1 2 2 3 7 7	(6)
(8)	9 8 7 5 3 2 2 2	22	1 3 5 6 6 8 9	(7)
(6)	8 7 6 5 2 1	23	4 5 7 8 8 9 9 9	(8)
(3)	8 6 3	24	2 4 5 6 7 8 8	(7)
(1)	0	25	0 2 7 8	(4)

Key: 5 | 22 | 6 means a reaction time of 0.225 seconds for A and 0.226 seconds for B

- (i) Find the median and the interquartile range for group A . [3]

The median value for group B is 0.235 seconds, the lower quartile is 0.217 seconds and the upper quartile is 0.245 seconds.

- (ii) Draw box-and-whisker plots for groups A and B on the grid. [3]



Q4 s2019_p62

- (i) Give one advantage and one disadvantage of using a box-and-whisker plot to represent a set of data. [2]

- (ii) The times in minutes taken to run a marathon were recorded for a group of 13 marathon runners and were found to be as follows.

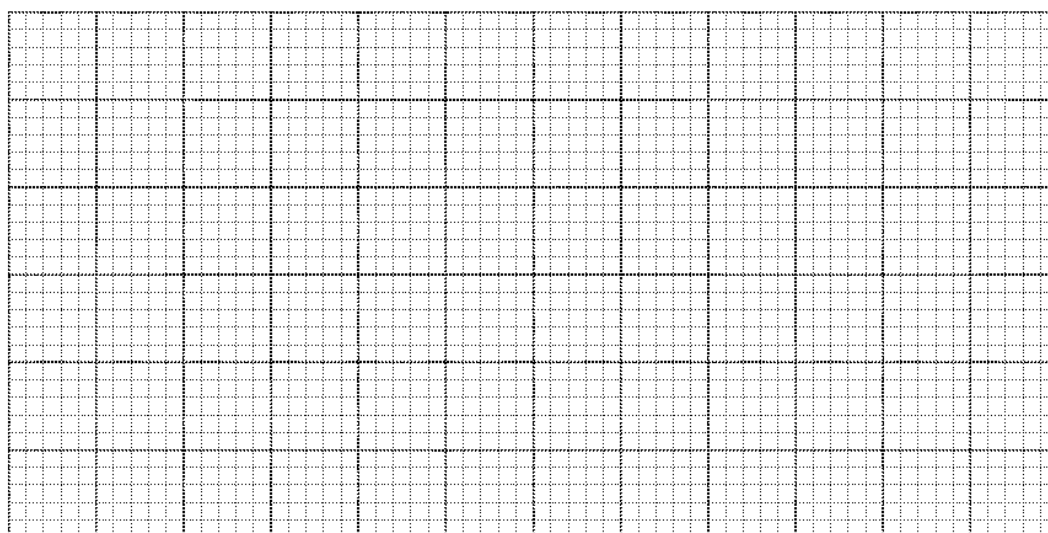
180 275 235 242 311 194 246 229 238 768 332 227 228

State which of the mean, mode or median is most suitable as a measure of central tendency for these times. Explain why the other measures are less suitable. [3]

- (iii) Another group of 33 people ran the same marathon and their times in minutes were as follows.

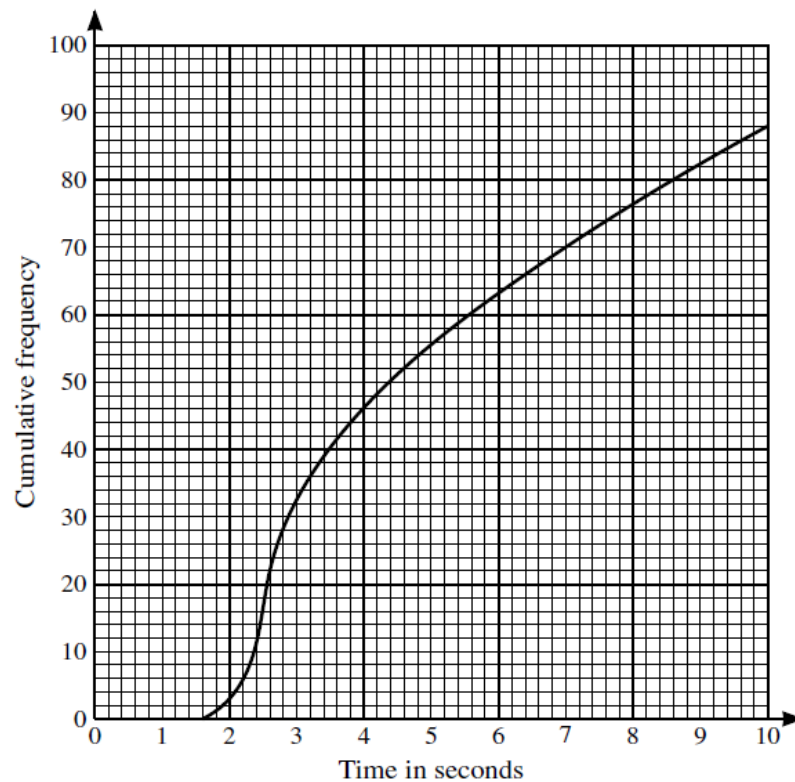
190 203 215 246 249 253 255 254 258 260 261
263 267 269 274 276 280 288 283 287 294 300
307 318 327 331 336 345 351 353 360 368 375

- (a) On the grid below, draw a box-and-whisker plot to illustrate the times for these 33 people. [4]



- (b) Find the interquartile range of these times. [1]

Q5 s2015_p62



In an open-plan office there are 88 computers. The times taken by these 88 computers to access a particular web page are represented in the cumulative frequency diagram.

- (i) On graph paper draw a box-and-whisker plot to summarise this information. [4]

An 'outlier' is defined as any data value which is more than 1.5 times the interquartile range above the upper quartile, or more than 1.5 times the interquartile range below the lower quartile.

- (ii) Show that there are no outliers. [2]

Q6 w2017_p63

The number of Olympic medals won in the 2012 Olympic Games by the top 27 countries is shown below.

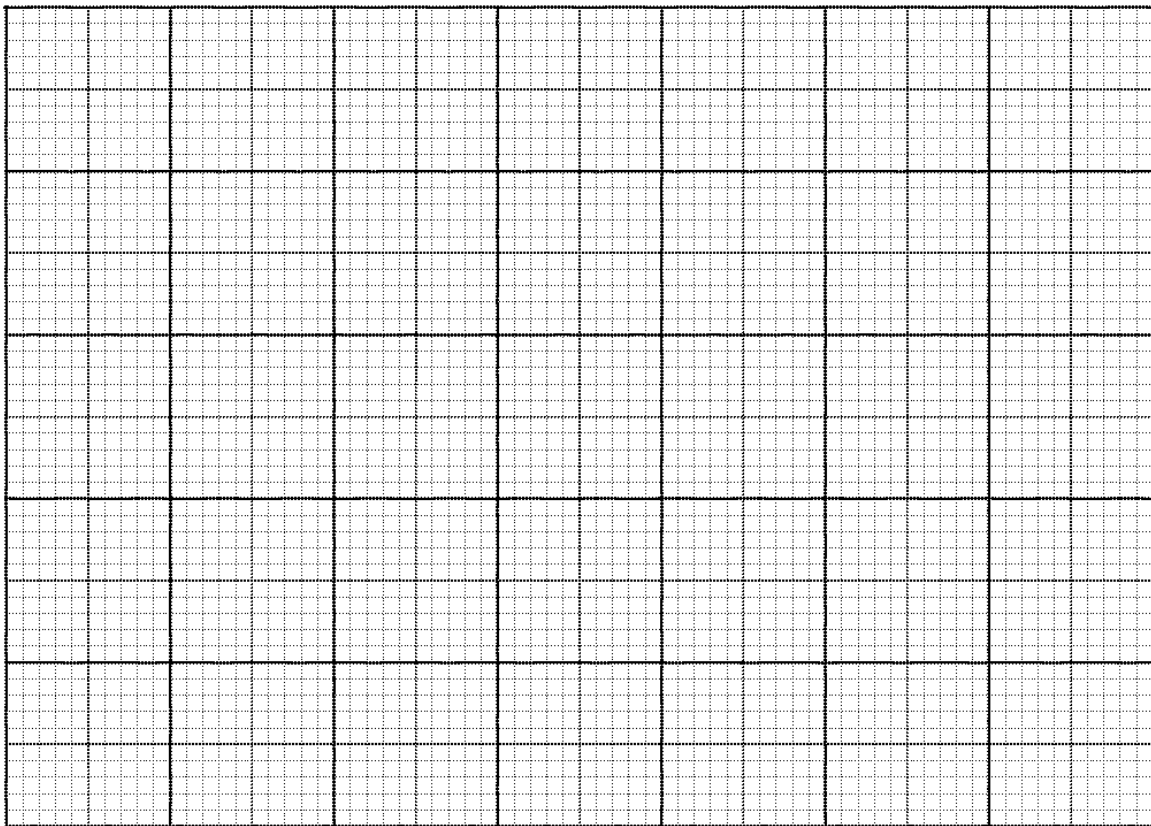
104	88	82	65	44	38	35	34	28
28	18	18	17	17	14	13	13	12
12	10	10	10	9	6	5	2	2

(i) Draw a stem-and-leaf diagram to illustrate the data.

[4]

(ii) Find the median and quartiles and draw a box-and-whisker plot on the grid.

[5]



ANSWERS

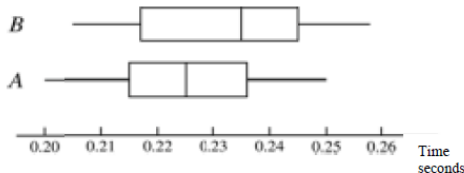
Q1

(i)		B1	Labels 'time' and 'seconds', 'boys' and 'girls' on correct plots and scaled line
		B1	One box and whisker all correct on graph paper – ignore boy or girl label
		B1	[3] Second box and whisker all correct (on graph paper and ignore boy/girl label) on SAME scaled line.
(ii)	girls smaller range or IQ range than boys /girls less spread out oe girls generally quicker than boys or girls median < boys median (not mean) oe boys almost symmetrical, girls +vely skewed oe	B1	Any 2 comments – MUST be a comparison
		B1	
		[2]	

Q2

(i)	LQ = 18, Median = 25, UQ = 50 	B1	median correct
		B1	LQ and UQ correct
		B1	Quartiles and median plotted as box graph with linear scale min 3 values
		B1ft	Whiskers drawn to correct end points with linear scale, not thr' box, not joining at top or bottom of box. Ft their UQ and LQ. Whiskers must be with ruler If scale non-linear or non-existent SCB1if all 5 data values (quartiles and end points) have values shown and all are correct numerically and fulfil the 'box' and 'whiskers ruled line' requirements
		B1	Label to include 'distance or travelled' and 'km,' allow 'total km', linear scale, numbered at least 5 – 70.
		5	
(ii)	1.5 × IQR = 48 Method 1 LQ – 48 = –ve, (i.e. < 0) UQ + 48 = 98 (i.e. > 70) hence no outliers	M1	Attempt to find 1.5 × their IQR and add to UQ or sub from LQ
		A1	Correct conclusion from correct working, need both ends. No need to state comparisons.
	Method 2 LQ – 5 = 13 (< 48) 70 – UQ = 20 (< 48) Hence no outliers	M1	Compare their 1.5 × IQR (= 48) > gap (20) between UQ and max 70 or LQ and min 5
		A1	Correct conclusion from correct working, need both ends. No need to state comparisons
		2	

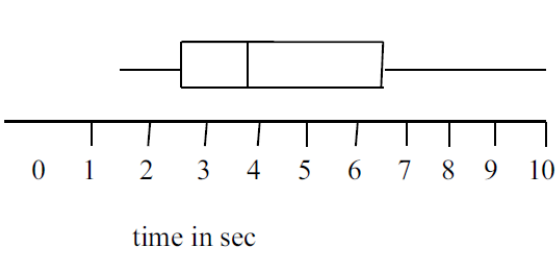
Q3

(i)	median = 0.225; LQ = 0.215: UQ = 0.236	B1	Correct median (Q_2)												
	IQR = 0.236 – 0.215	M1	$0.232 < \text{UQ } (Q_3) < 0.238 - 0.204 < \text{LQ } (Q_1) < 0.219$												
	= 0.021	A1	www Omission of all decimal points MR-1 <u>If M0 awarded</u> SCB1 for both LQ = 0.215: UQ = 0.236 seen												
		3													
(ii)		B1	Linear scale between 0.20 to 0.26 (condone omission of 0.26) axis labelled (time and seconds), at least one box plot attempted, no lines through boxes, whiskers not at corner of boxes												
		B1 ft	Labelled correct graph for A, (ft their median/quartiles), condone lines through boxes, whiskers at corner of boxes												
	<table border="1"><tr><td>A</td><td>0.200</td><td>0.215</td><td>0.225</td><td>0.236</td><td>0.250</td></tr><tr><td>B</td><td>0.205</td><td>0.217</td><td>0.235</td><td>0.245</td><td>0.258</td></tr></table>	A	0.200	0.215	0.225	0.236	0.250	B	0.205	0.217	0.235	0.245	0.258	B1	Labelled correct graph for B, condone lines through boxes, whiskers at corner of boxes SC If B0B0 scored because graphs not labelled/labels reversed SCB1 if both 'correct' Penalty MR-1 if graphs plotted on separate axes unless both scales align exactly.
	A	0.200	0.215	0.225	0.236	0.250									
B	0.205	0.217	0.235	0.245	0.258										
	3														

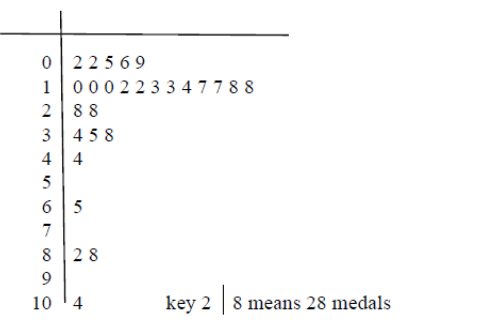
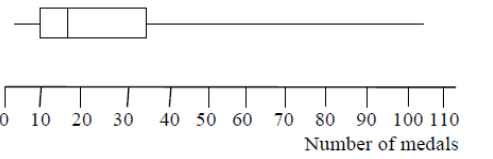
Q4

(i)	Advantage: comment referring to spread or range or shape	B1	Comments referring to quartiles, IQR, Range, median, shape, skewness, data distribution, spread score B1 Any comments with reference to mean or standard deviation or any other 'disadvantage' will score B0 Comments referring to '5-value plot', comparison with another data set, overview or ease of drawing/plotting/reading require an appropriate advantage statement.
	Disadvantage: comment referring to limited data information provided	B1	Comments referring to no individual data, no information about the number of values, unable to calculate mean, standard deviation, variance and mode score B1 Any comments with reference to median, shape or any other 'advantage' will score B0 Comments referring to 'size of data set' or 'average' require an appropriate disadvantage statement. Comments referring to outliers are ignored in all cases (as outliers are not in the syllabus content) unless supported by an appropriate advantage / disadvantage statement. If comments not clearly identified, assume first comment is the advantage.
		2	
(ii)	Not mean as data skewed by one large value	B1	Comment which identifies 768 (or 'a very large number') as the problem. Condone the use of 'outlier'
	Not mode as frequencies all the same	B1	Comment which indicates that no mode exists (e.g. all the data is different, there is no repeated number, all the values are different)
	Median	B1	Median identified as choice, dependent upon statements for mean and mode being given, even if incorrect or very general.
	SC: Mean is identified as most suitable		
	Not mode as frequencies all the same	SCB1	Comment which indicates that no mode exists
	Not median as not all values used	SCB1	Comment which indicates limitation of median e.g. median is not in middle of range.
		3	
(iii)(a)	LQ = 256 or 256.5 Med = 280 UQ = 329 Min 190 max 375	B1	Median, UQ and LQ values seen, may not be identified or identified correctly. (Not read from box plot unless value stated)
		B1	FT Median and quartiles plotted in box on graph, linear scale
		B1	Correct end points, whiskers from ends of box but not through box, not at top or bottom of box
		B1	Uniform scale from 190 to 375 (need at least 3 linear identified points min) and labelled 'time' and 'minutes' (can be in title) No time axis or time axis with no scale attempt, Max B1B0B0B0
		4	
(iii)(b)	IQR = <i>their</i> 329 – <i>their</i> 256 = 73 or 72.5	B1	FT Must follow through only from <i>their</i> stated values (condone if correct quartiles stated here), not reading from graph.
		1	

Q5

(i)		B1	LQ = 2.6 med = 3.8– 3.85, UQ = 6.4– 6.6
		B1 ✓	Correct quartiles and median on graph fit linear from 2–10
		B1	End whiskers correct not through box
		B1	4 Label need seconds and linear 2–10 axis or can have 5 values on boxplot no line provided correct
(ii)	$1.5 \times \text{IQR} = 1.5 \times 3.8 = 5.7$ $\text{LQ} - 5.7 = -\text{ve}, \text{UQ} + 5.7 = 12.1 \text{ i.e. } > 10$ So no outliers AG	M1	Attempt to find $1.5 \times \text{IQR}$ and add to UQ or sub from LQ OR compare $1.5 \times \text{IQR}$ with gap 3.6 between UQ and max 10
		A1	2 Correct conclusion from correct working need both

Q6

(i)		B1	Stem, digits 5, 7, 9 can be missing here, can be upside down
		B1	All leaves in correct order increasing from stem, (5, 7 and 9 can be missing), condone commas
		B1	Reasonable shape, requires all values of the stem, only one line for each stem and leaves must be lined up. Can be upside down or sideways. No commas. Condone one 'leaf' error.
		B1	Correct key must state 'medals' or have 'medals' in leaf heading or title
		4	
(ii)	Med = 17 LQ = 10 UQ = 35 	B1	Median correct
		B1	LQ and UQ correct
		B1	Uniform scale from 2 to 104 (need 3 identified points min) and label including medals (can be in title)
		B1 FT	Correct box med and quartiles on diagram, FT their values
		B1	Correct end-whiskers from ends of box but not through box
		5	